Bot Attacks: Top Threats and Trends

Vol. 1 September 2021

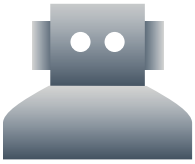
Insights into the growing number of automated attacks

Not all bots are created equal. While some bots, such as search engine crawlers, are good, bad bots are built to carry out malicious attacks at scale. Traffic from these bad bots is exploding, and this in-depth report explores emerging traffic patterns, live examples of bot behavior and detection, and the steps you should take to protect your business.»

Barracuda
Your journey, secured.

Table of Contents

Key findings



Bots make up nearly **two-thirds** of internet traffic

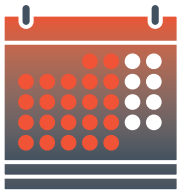
E-commerce applications and **login portals** are the most common targets of **advanced persistant bots**

North America accounts for **67%** of bad bot traffic

aws

Microsoft Azure

Most bot traffic comes in from **two** large **public clouds: AWS** and **Microsoft Azure**



Bad bots follow a standard **workday**

Introduction

Over the past few years, automated bot traffic has grown rapidly. Once used primarily by search engines, bots now have a variety of uses — both good and bad. The good bots are primarily search engine crawlers, social network bots, aggregator crawlers, monitoring bots, etc. These bots obey the website owner's rules as specified in the robots.txt file, publish methods of validating them as who they say they are, and work in a way to avoid overwhelming the websites and applications they visit.

Bad bots are built to perform various malicious activities. They range from basic scrapers that try to get some data off an application (and are easily blocked) to advanced persistent bots that behave almost like human beings and look to evade detection as much as possible. These bots attempt attacks such as web and price scraping, inventory hoarding, account takeover attacks, distributed denial of service (DDoS) attacks, and much more. Bad bots make up a significant part of website traffic today, and detecting and blocking them is of critical importance to businesses.

Barracuda researchers have analyzed the traffic patterns measured by Barracuda application security solutions over the first six months of 2021, and in this report we'll share the insights they uncovered both in terms of traffic trends and live examples of bot behavior and detection.

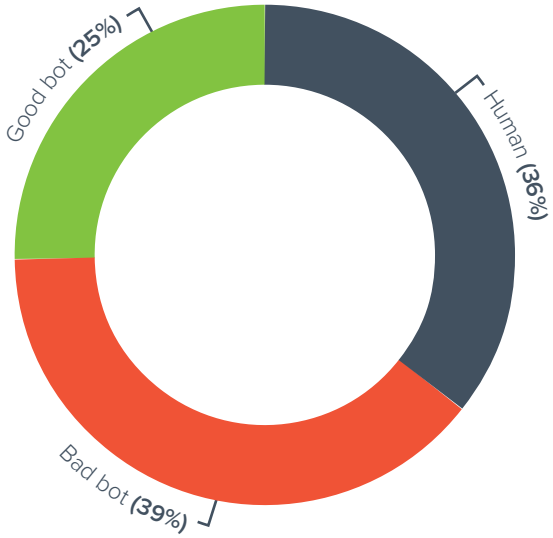
Traffic Trends

Insight 1: Bots make up 64% of internet traffic

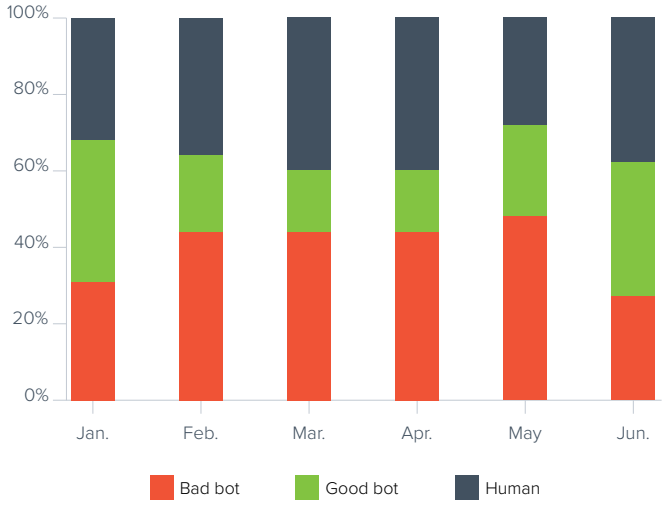
Automated traffic makes up nearly two-thirds of internet traffic, as measured by Barracuda technology over the first six months of 2021. Roughly 25% of this traffic is from known good bots — ones like search engine crawlers, social network bots, and monitoring bots.

However, our measurements show that nearly 40% of traffic in total was from bad bots. These bad bots include both basic web scrapers and attack scripts, as well as advanced persistent bots. These advanced bots try their best to evade standard defenses and attempt to perform their malicious activities under the radar. In our dataset, the most common of these persistent bots were ones that went after e-commerce applications and login portals.

Traffic distribution: Bots vs. humans (January – June 2021)



Distribution by month



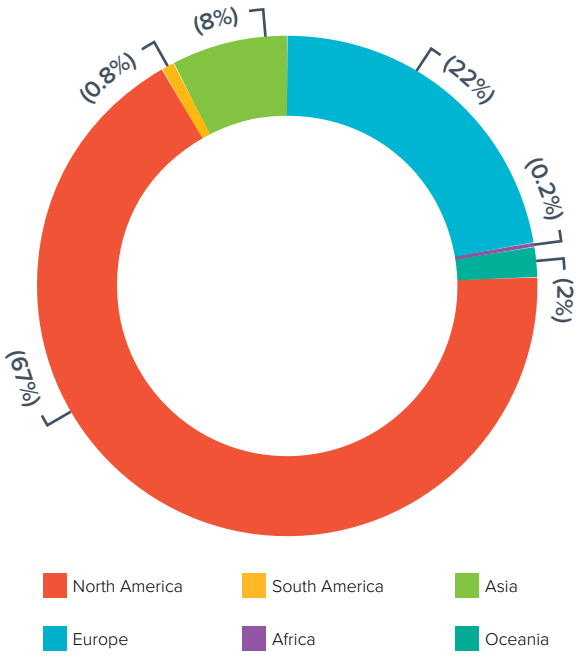
Insight 2: North America accounts for the largest portion of bad bot traffic — and most of it originates from data centers

Most of the bad bot traffic comes in from data centers' IP ranges. This makes it relatively simple to identify and block these bots. If your application does not expect traffic from a specific data center IP range, you can consider blocking it, similar to geo-IP based blocking.

From our sample set, most of the bot traffic was coming in from the two large public clouds — AWS and Microsoft Azure — in roughly equal measure. This could be because it is easy to set up an account for free with either provider and then use the account to set up the bad bots.

Looking at regional traffic distributions, North America accounts for 67% of bad bot traffic, followed by Europe and then Asia. Interestingly, the European bot traffic is more likely to come in from hosting services (VPS) or residential IPs than the North American traffic.

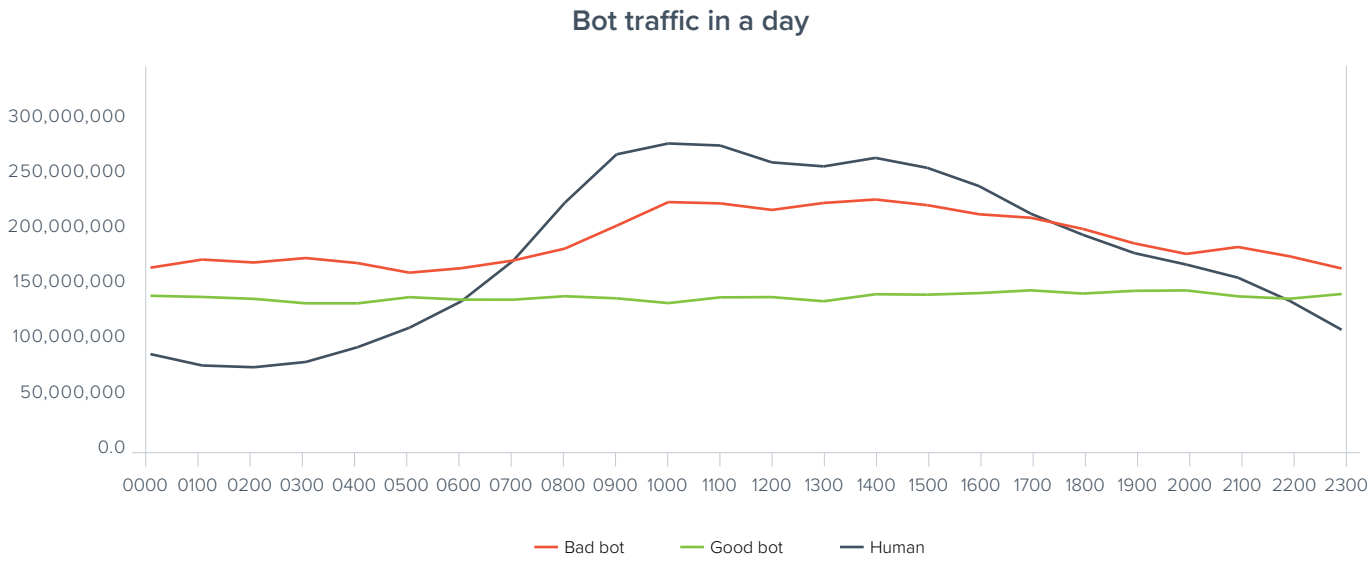
Geographical sources of bad bot traffic



Insight 3: Bad bots follow a standard workday

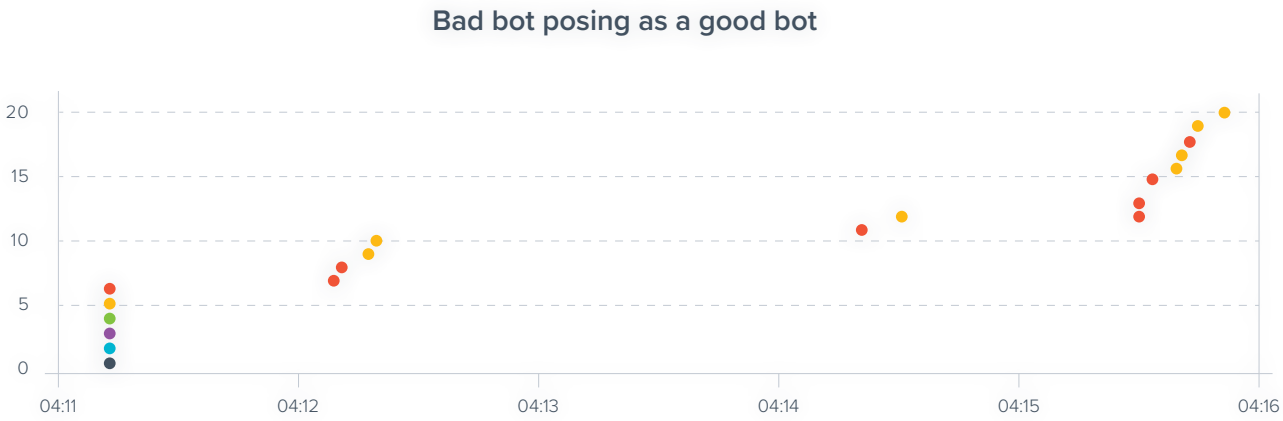
In 2020, our researchers found that bad bot traffic typically follows the standard workday. Our analysis for the first half of 2021 confirms this. Good bots follow a normal distribution — they don't vary much, and the traffic rate is fairly constant through the day. In the six months we analyzed, a good chunk of this traffic is from monitoring bots, and this lack of variance is expected.

However, when it comes to bad bots, they follow the standard workday — and with good reason. The attackers running these bad bots prefer to hide within the normal human traffic stream to avoid raising alarm bells. The common stereotype of a "hacker" performing their attacks late into the night in a dark room with green fonts on a black screen has been replaced by people who set up their bots to carry out the automated attacks while they go about their day.



Real-life examples of bad bots

Example 1: Pretending to be a known vulnerability scanner



Each dot represents a unique URL being accessed. The first group shows that a lot of requests to any URL were made at the same time, and at this point the client was stopped. After the first burst, the requests were coming in smaller amounts but only to specific URLs, and the bot was identified using its characteristics.

In our analysis, we found this example of a bad bot pretending to be a known vulnerability scanner (a good bot). The bad bot was attempting to perform reconnaissance and probe for vulnerabilities using some basic attacks. As such, the bot was using a standard browser user agent, but it had additional custom HTTP headers that spoofed the headers of a scanner used by the organization being attacked.

The bot, however, failed on multiple counts and was caught relatively easily. One telltale sign was that the fingerprint of the client did not match that of a known browser. While the custom headers being spoofed were correct, the header order being sent by the tool did not fit the expected profile. The bad bot, which came in from residential IP addresses, was also visiting pages at random. All these actions together were used to detect it and block its persistent attempts quite quickly.

Example 2: Accessing the login page of a medical service provider

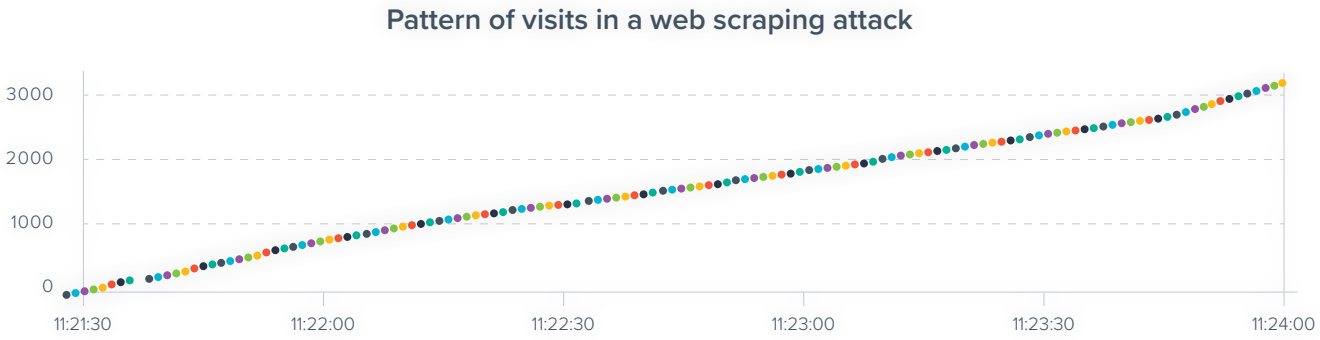
In this example that we detected, the bot was accessing the login page of a medical service provider. Pretending to be Internet Explorer on Windows 10, this bot was also appending random UTM parameters to the login page URL and coming in from multiple AWS IP ranges.

The bot was detected based on the header variations that differed from the headers of a standard browser. It also stood out due to the fact that the same browser signature was coming in from multiple AWS IP addresses but accessing only the login pages on the application. The giveaway was a brute force attempt using stolen credentials. This was caught using our credential database, and the bot was blocked from accessing the site.

This graph shows a sample of the login requests performed by the bot while attempting to brute force the login page. Each dot represents a login attempt.

Example 3: Web scraping a B2B e-commerce store

This bot was caught attempting to scrape a lot of information from a business-to-business (B2B) e-commerce store in the UK. The bot was coming in as a standard browser and had all its headers in order. It was also coming in from a residential IP address, which is where the system got its first hint of a problem — this website was very rarely accessed by residential customers. In addition, the client was detected to be using a Web SDK kit, typically used for automation, and these detections, along with the rapid traversal of the website, were used to detect and block the bot.

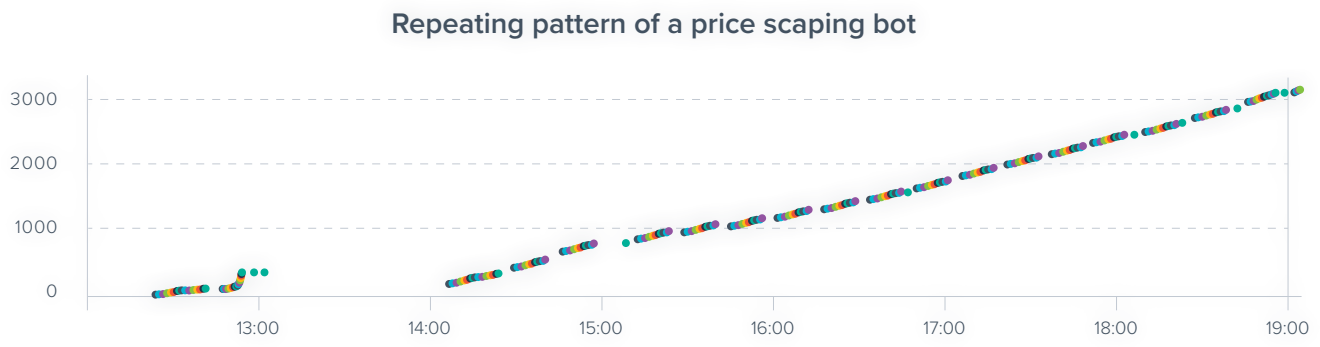


This graph shows the bot attempting to access the same set of URLs multiple times within a short period of time to scrape data. It would perform this pattern multiple times during the day.

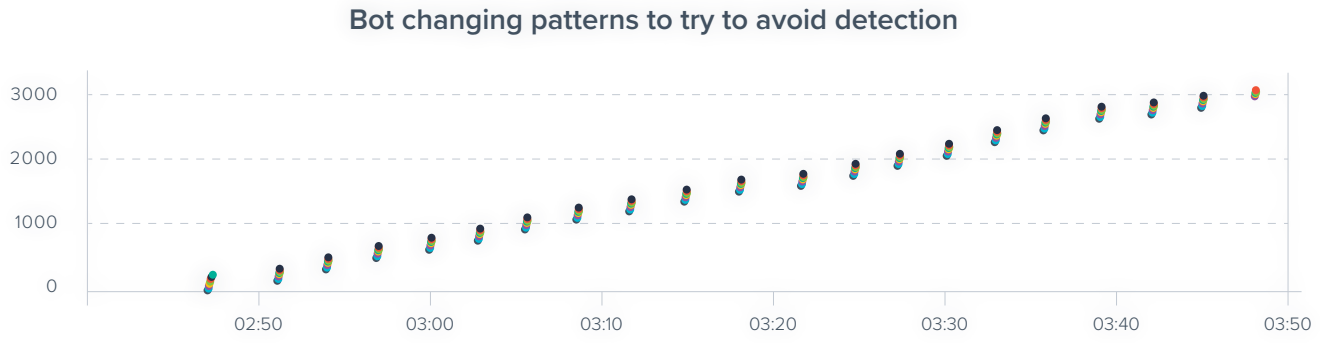
Example 4: Price scraping an e-commerce store in Eastern Europe

In this example, there was a suspected price scraping attempt on an e-commerce store based in Eastern Europe. The store was running a discount on Apple products, and there were some suspicious patterns of behavior in the traffic. The suspicious traffic came with standard browser clients, through multiple local residential IP addresses. However, these local IP addresses were from VPS hosting providers, and each client would only access a standard set of pages. After a few iterations of these requests, the traffic patterns were correlated, identified as price scraping attempts, and blocked.

Once these scrapers were blocked, the bad bots started coming in with different browser patterns and from IP addresses in neighboring countries. The activity was easier to identify this time because the new clients used non-standard browser headers and they were accessing the same pages and also got blocked.



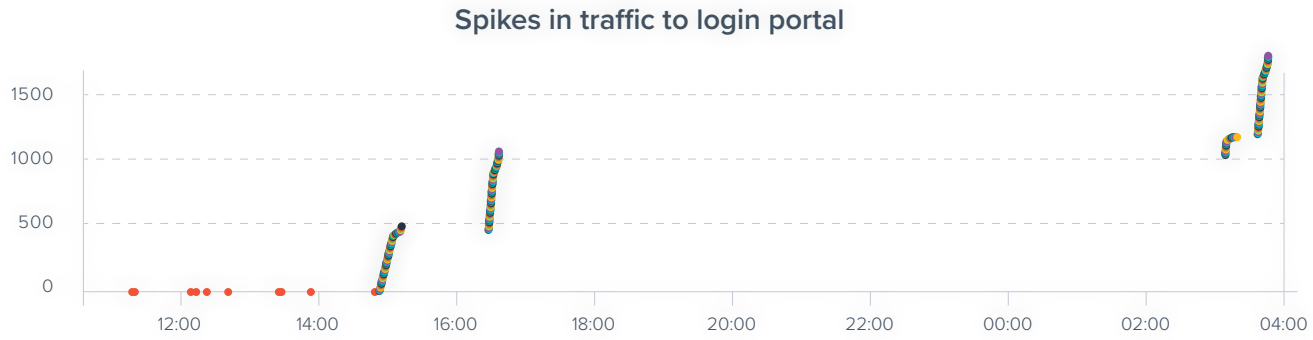
Bots were accessing the same set of product URLs multiple times in an hour after the initial burst was blocked.



Bots attempting to access a smaller set of product pages in a different browsing pattern multiple times an hour.

Example 5: Attempting to overwhelm the login portal of an Indian manufacturing company

In this example, the login portal of an Indian manufacturing company was seeing unusually high traffic. The traffic was coming in primarily from mobile networks, which was unusual, but not unexpected for this website. However, on further analysis, the system determined that the incoming traffic was more likely from a desktop browser that was impersonating a mobile device while connected to a hotspot. The multiple clients attempting to overwhelm this login page were blocked successfully, and the page response time came back to normal.



The first few dots were a bot pretending to be human and spreading out its accesses. After that, there are clusters seen, and each dot represents a different client attempting to access the login page.

Best practices to protect against bot attacks

Bad bots are a big problem for web and API application owners today. These malicious bots attack user accounts, skew analytics, scrape data, and destroy customer experience. And ultimately, they can lead to a data breach. According to The State of Application Security in 2021, bot-based attacks are the most likely contributor to successful security breaches resulting from application vulnerabilities in the past 12 months.

When it comes to protecting against newer attacks, such as bots, defenders can be overwhelmed at times due to the number of solutions required. The good news is that solutions are consolidating into WAF/WAF-as-a-Service offerings, also known as Web Application and API Protection (WAAP) services.

To protect your business, as well as your data, analytics, and inventory, you need to invest in WAAP technology that identifies and stops bad bots in their tracks. This will improve both user experience and overall security.

- **Put proper application security in place.** Install a web application firewall or WAF-as-a-Service solution and make sure it is properly configured. This is an important first step to make sure your application security solution is working as intended.

- **Invest in bot protection.** Make sure the application security solution you choose includes anti-bot protection so it can effectively detect and stop advanced automated attacks.

- **Take advantage of machine learning.** With a solution that uses the power of machine learning, you can effectively detect and block hidden almost-human bot attacks. Be sure to turn on credential stuffing protection to prevent account takeover as well.

About Barracuda

At Barracuda, we strive to make the world a safer place.

We believe every business deserves access to cloud-enabled, enterprise-grade security solutions that are easy to buy, deploy and use. We protect email, networks, data and applications with innovative solutions that grow and adapt with our customers' journey.

More than 200,000 organizations worldwide trust Barracuda to protect them — in ways they may not even know they are at risk — so they can focus on taking their business to the next level.

For more information, visit barracuda.com.

Barracuda
Your journey, secured.